STIC Search Report

EIC 2100

USPTO

TO: Kuen S Lu **Location: RND 3B02** **Art Unit: 2167** **Monday, January 09, 2006** **Case Serial Number: 09/885902**	**From: Ruth E. Spink** **Location: EIC 2100** **RND-4B31** **Phone: 23524** **Ruth.spink@uspto.gov**

Search Notes

Kuen – Attached is the foreign patent and NPL search for the above referenced case. I tagged a few that I thought might be of particular interest. Be sure to let me know if you would like for me to refocus the search.

Ruth

STIC EIC 2100
Search Request Form 175917

| Today's Date: 1/9/2005 | What date would you like to use to limit the search? |
| | Priority Date: 6/21/2001   Other: |

Name Kuen S. Lu

AU 2167   Examiner # 7999/
571-272-4114

Room # RAN 3B02   Phone 24114

Serial # 09/885, 902

Format for Search Results (Circle One):

(PAPER)   DISK   EMAIL

Where have you searched so far?   Citeseer

(USP)   DWPI   (EPO)   (JPO)   (ACM)   IBM TDB Sun

IEEE   INSPEC   SPI   Other google, Oracle

Is this a "Fast & Focused" Search Request? (Circle One) (YES)   NO
A "Fast & Focused" Search is completed in 2-3 hours (maximum). The search must be on a very specific topic and meet certain criteria. The criteria are posted in EIC2100 and on the EIC2100 NPL Web Page at http://ptoweb/patents/stic/stic-tc2100.htm.
Please complete by : Am 1/05 pm THANKS

What is the topic, novelty, motivation, utility, or other specific details defining the desired focus of this search? Please include the concepts, synonyms, keywords, acronyms, definitions, strategies, and anything else that helps to describe the topic. Please attach a copy of the abstract, background, brief summary, pertinent claims and any citations of relevant art you have found.

As attached claims

STIC Searcher _____   Phone _____

Date picked up _____   Date Completed_____

gira

STIC Search Results
Feedback Form

EIC 2100

Questions about the scope or the results of the search? Contact *the EIC searcher or contact*:

Anne Hendrickson, EIC 2100 Team Leader
272-3490, RND 4B28

Voluntary Results Feedback Form

➤ *I am an examiner in Workgroup:* [   ]   *Example: 2133*

➤ *Relevant prior art **found**, search results used as follows:*

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
  (journal articles, conference proceedings, new product announcements etc.)

➤ *Relevant prior art **not found**:*

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/EIC2100 RND, 4B28

```
Set      Items    Description
S1      412774    SEARCH?? OR SEARCHING OR FIND? ? OR FINDING OR FOUND OR RE-
                  TRIEVE? ? OR RETRIEVING OR RETRIEVAL OR QUERY OR QUERIES OR Q-
                  UERYING
S2      310187    KEYWORD? ? OR PHRASE? ? OR TERM? ? OR WORD? ?
S3     1714511    ATTRIBUTE? ? OR CHARACTERISTIC? ? OR PROPERTY OR PROPERTIES
                  OR METADATA
S4     1044461    SCALE? ? OR SCALING OR SCORE? ? OR SCORING OR WEIGHT?? OR -
                  WEIGHTING
S5      749917    ORDER?? OR ORDERING OR SORT?? OR SORTING
S6          40    S1 AND S2 AND S3 AND S4 AND S5
S7          23    S6 AND IC=G06F
S8          23    IDPAT (sorted in duplicate/non-duplicate order)
S9          23    IDPAT (primary/non-duplicate records only)
File 347:JAPIO Nov 1976-2005/Aug(Updated 051205)
         (c) 2005 JPO & JAPIO
File 350:Derwent WPIX 1963-2006/UD,UM &UP=200602
         (c) 2006  Thomson Derwent
```

017433665
WPI Acc No: 2005-757344/200577
Related WPI Acc No: 2006-020070
XRAM Acc No: C05-231131
XRPX Acc No: N05-624936
**Procuring biological content on electronic file, by interfacing user with
server that accesses medium with files of targets, inputting request to
generate biological** attribute **extracts, generating hierarchical output
based on extracts**
Patent Assignee: INVITROGEN CORP (INVI-N)
Inventor: LIANG F
Number of Countries: 001 Number of Patents: 001
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
US 20050240352	A1	20051027	US 2004830074	A	20040423	200577	B

Priority Applications (No Type Date): US 2004830074 A 20040423
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
US 20050240352	A1		36	G06F-019/00	

Abstract (Basic): US 20050240352 A1
        NOVELTY - Procuring (M1) biological content and their
products/services listed on electronic inventory file comprising
interfacing by user through user terminals and bi-directional
communication connections with target item server which accesses
electronic storage medium having files comprising grouping of target
items, inputting request to generate extracts of biological **attribute**
, generating page having hierarchical menu output based on extracts, is
new.
        DETAILED DESCRIPTION - Procuring (M1) biological content and their
products and/or services listed on an electronic inventory file, where
the inventory file is stored on one or more electronic storage medium
which comprises several files comprising one or more segregated sundry
grouping of target items, involves:
        (a) interfacing by one or more user through user terminals and
bi-directional communication connections with one or more target item
server which accesses the electronic storage medium, where extracts
comprising one or more associated biological **attribute** are generated
in the server for the target items in the electronic storage medium
through an appropriate request;
        (b) inputting a request to generate the extracts;
        (c) **retrieving** the extracts; and
        (d) generating a page comprising one or more hierarchical menu
output based on such extracts that provides one or more user, one or
more subset of the target items stored on the electronic medium, where
the one or more menu **sorts** the target items in the subset into a user
accessible file of target items based on a empirical measure of
similarity of the associated biological **attributes** for the **sorted**
target items, and where the one or more hierarchical menu output
display page identifies the target items **sorted** into each file which
have one or more associated biological **attribute** in common to enable
one or more user to differentiate products and/or services of interest
stored on the electronic storage medium and to procure the differential
products by activating an appropriate graphic user interface (GUI)
comprising the displayed output page.
        INDEPENDENT CLAIMS are also included for the following:
        (1) a server (I) configuration for carrying out (M1); and
        (2) offering (M2) a product or service to a user in a remote
location, involves remotely providing an electronic data server to the
user, receiving an input from the user, processing the input to produce

a first output, interfacing one or more public consortium database with
one or more database proprietary to an offerer of the product or
service, selecting a first product or service or a link or description
of a first product or service to create an extract, and outputting the
extract to the user.

USE - (M1) or (I) is useful for procuring biological content and
their products and/or services listed on an electronic inventory file.
The products and/or services are biologically related products and/or
services, where the biologically related products are chosen from
cloned nucleic acid inserts comprising a structural gene or
transcriptional unit, bioassays, labeling and detection dyes, vectors,
antibodies, peptides, nucleic acids, enzymes, nucleotides, buffers,
cells media, selection molecules, expression systems, lipids,
transfection reagents, electrophoresis products, separation column,
affinity compounds, membranes, open reading frames (ORFs), DNA and RNA
primers and proteins (claimed).

ADVANTAGE - (M1) efficiently **searches** and extracts relevant data.
(M1) enables linking of biological information to E-commerce through
effective information browsing, processing and reporting.

pp; 36 DwgNo 0/16

Title Terms: BIOLOGICAL; CONTENT; ELECTRONIC; FILE; INTERFACE; USER; SERVE;
  ACCESS; MEDIUM; FILE; TARGET; INPUT; REQUEST; GENERATE; BIOLOGICAL;
  ATTRIBUTE ; EXTRACT; GENERATE; HIERARCHY; OUTPUT; BASED; EXTRACT
Derwent Class: B04; D16; S05; T01
International Patent Class (Main): **G06F-019/00**
International Patent Class (Additional): G01N-033/48; G01N-033/50
File Segment: CPI; EPI

017135672
WPI Acc No: 2005-460017/200547
XRPX Acc No: N05-373724
  Method and system of public communication source guiding
Patent Assignee: JIAWA SCI & TECHNOLOGY CO LTD NANJING (JIAW-N)
Inventor: HUANG H; LU S; ZHOU H
Number of Countries: 001 Number of Patents: 001
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
CN 1598489	A	20050323	CN 200441954	A	20040914	200547	B

Priority Applications (No Type Date): CN 200441954 A 20040914
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
CN 1598489	A			G01C-021/26	

Abstract (Basic): CN 1598489 A
      NOVELTY - The invention is a method for public resource navigation,
based on the computer system for public traffic resources databank, it
designs data structure storage and **searching** data joint, the storing
mode of data structure for point in databank is mainly stored in two
data tables, namely the point basic **attributes** table and information
table of relative station.
      DETAILED DESCRIPTION - The **attributes** stored in the point basic
information table has: point name, **searching** key **word** , point sketch
map, type and other basic **attributes** ; the joint is bonded with point
data, and arranges with **ordering** joint, the **ordering** joint is made
up of three parts: the number of this joint, the **weight** of the
distance between the two joints from this joint to the next joint and
the next joint; acquires all the relative information from the two data
tables, and they are displayed on the web.
      DwgNo 0/1
Title Terms: METHOD; SYSTEM; PUBLIC; COMMUNICATE; SOURCE; GUIDE
Derwent Class: T01
International Patent Class (Main): G01C-021/26
International Patent Class (Additional): **G06F-017/30** ; G08G-001/00
File Segment: EPI

016430355   **Image available**
WPI Acc No: 2004-588272/200457
XRPX Acc No: N04-465590
   Document retrieval **device calculates** score **of document by multiplying number of lexicon with output of collation unit which collates** keyword **output from user's terminal with** keyword **received by lexicon, load coefficient**
Patent Assignee: MITSUBISHI ELECTRIC CORP (MITQ)
Number of Countries: 001 Number of Patents: 001
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
JP 2004234516	A	20040819	JP 200324524	A	20030131	200457	B

Priority Applications (No Type Date): JP 200324524 A 20030131
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 2004234516	A		12	G06F-017/30	

Abstract (Basic): JP 2004234516 A
        NOVELTY - An evaluation unit (19) calculates **score** of the
    document by multiplying the number of lexicon with the output of a
    collation unit (18) which collates the **keyword** output from user's
    terminal with **keyword** received by the lexicon, load coefficient
    produced by comparing various document with respect to recording
    content. An **order** form is presented to the user based on the
    calculated **score** .
        USE - For **retrieval** of document using personal computer (PC),
    personal digital assistant (PDA), mobile telephone.
        ADVANTAGE - **Searches** desired document effectively, without
    performing special setting operation.
        DESCRIPTION OF DRAWING(S) - The figure shows the block diagram of
    the document **retrieval** device. (Drawing includes non-English language
    text).
        load coefficient determination unit (15)
        document **characteristic** analysis unit (16)
        document **characteristics** storage (17)
        collation unit (18)
        document evaluation unit (19)
        pp; 12 DwgNo 1/12
Title Terms: DOCUMENT; **RETRIEVAL** ; DEVICE; CALCULATE; **SCORE** ; DOCUMENT;
   MULTIPLICATION; NUMBER; OUTPUT; COLLATE; UNIT; COLLATE; **KEYWORD** ; OUTPUT
   ; USER; TERMINAL; **KEYWORD** ; RECEIVE; LOAD; COEFFICIENT
Derwent Class: T01
International Patent Class (Main): **G06F-017/30**
File Segment: EPI

9/5/6   (Item 6 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2006 Thomson Derwent. All rts. reserv.

015580913   **Image available**
WPI Acc No: 2003-643070/200361
   Searching **method and system based on example for deciding similarity**
Patent Assignee: POSCO (POSC-N); UNIV POHANG SCI & TECHNOLOGY (UYPO-N);
   POSTECH FOUND (POST-N)
Inventor: KIM J S; KWON O U; LEE J H; PARK J S; PI Y J; SONG N G
Number of Countries: 002 Number of Patents: 002
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
KR 2003039576	A	20030522	KR 200170541	A	20011113	200361	B
JP 2003281186	A	20031003	JP 2002322059	A	20021106	200367	

Priority Applications (No Type Date): KR 200170541 A 20011113
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
KR 2003039576	A		1	G06F-017/30	
JP 2003281186	A		14	G06F-017/30	

Abstract (Basic): KR 2003039576 A
     NOVELTY - A **searching** method and system based on an example for
deciding similarity is provided to rapidly and exactly determine
identification and similarity of related techniques.
     DETAILED DESCRIPTION - A related technique document is inputted in
a related document input unit(311) of an index unit(310). A paragraph
is divided by a structural **characteristic** of the inputted related
technique document, and a **keyword** is extracted in the first
extracting unit(312) according to divided paragraphs. A **weight** value
in each paragraph with respect to the extracted **keyword** is obtained,
and a **keyword** and a **weight** value thereof are expressed as a unit
vector in the first **word** vector expression unit(313). The **keyword**
and a **weight** value expressed as a unit vector are stored in a unit
vector storing unit(314). An example document having an example
technique is inputted in an example document input unit(321). A
paragraph is divided in accordance with a structural **characteristic**
in the inputted example document, and a **keyword** is extracted in the
second **keyword** extracting unit(322) according to divided paragraphs.
A **weight** value is obtained in each paragraph, and a **keyword** and a
weight value thereof are expressed as a unit vector in the second
word vector expression unit(323). A similarity calculation unit(324)
obtains a similarity between corresponding paragraphs with an example
document and a related technique document, and obtains a similarity
between the example document and a related technique document using a
similarity between paragraphs. A display unit(325) **sorts** related
technique documents in ascending powers of the obtained similarity and
supplies the documents for a user.
     pp; 1 DwgNo 1/10
Title Terms: **SEARCH** ; METHOD; SYSTEM; BASED; EXAMPLE; DECIDE; SIMILAR
Derwent Class: T01
International Patent Class (Main): **G06F-017/30**
File Segment: EPI

013765703   **Image available**
WPI Acc No: 2001-249914/200126
XRPX Acc No: N01-178241
   **Dictionary production assistance apparatus for use during processing of
   Japane**se language documents, evaluates characteristics **of each bigram
   using evaluation** scale **representing degree of importance**
Patent Assignee: HITACHI LTD (HITA   )
Number of Countries: 001 Number of Patents: 001
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
JP 2001043226	A	20010216	JP 99219562	A	19990803	200126	B

Priority Applications (No Type Date): JP 99219562 A 19990803
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 2001043226	A		10	G06F-017/28	

Abstract (Basic): JP 2001043226 A
     NOVELTY - A collecting unit (1A) collects bigrams present in input
document data (1011) and counter counts the number of occurrence of
bigrams in collector data. An evaluating unit evaluates
characteristics of each bigram using evaluation **scale** representing
degree of importance. A display unit (106) displays evaluated bigrams
that satisfy predefined conditions during evaluation in **order** of
degree of importance.
     DETAILED DESCRIPTION - An INDEPENDENT CLAIM is also included for
recording medium.
     USE - For dictionary production assistance for use during
processing of Japanese language documents, for production of index
vocabulary in information **retrieval** , and for production of machine
translation dictionary.
     ADVANTAGE - By producing dictionary based on degree of importance
of bigrams, common bigrams present in document data with high degree of
coincidence are removed automatically, hence production efficiency of
dictionary with essential **words** is enhanced.
     DESCRIPTION OF DRAWING(S) - The figure shows the block diagram of
word dictionary production assistance apparatus. (Drawing includes
non-English language text).
     Collecting unit (1A)
     Display unit (106)
     Input document data (1011)
     pp; 10 DwgNo 1/9
Title Terms: DICTIONARY; PRODUCE; ASSIST; APPARATUS; PROCESS; JAPAN;
   LANGUAGE; DOCUMENT; EVALUATE; **CHARACTERISTIC** ; EVALUATE; **SCALE** ;
   REPRESENT; DEGREE; IMPORTANT
Derwent Class: P86; T01; W04
International Patent Class (Main): **G06F-017/28**
International Patent Class (Additional): **G06F-017/22** ; **G06F-017/27** ;
   G06F-017/30 ; G10L-015/18
File Segment: EPI; EngPI

DIALOG(R)File 350:Derwent WPIX
(c) 2006 Thomson Derwent. All rts. reserv.

013251186   **Image available**
WPI Acc No: 2000-423069/200036
XRPX Acc No: N00-315724
  Characterizing term **extraction method in computer, involves** sorting
  extracted terms **according to generated moduli and accepting** terms
  with greatest moduli as characteristic keyword **of documents content**
Patent Assignee: JUSTSYSTEM PITTSBURGH RES CENT INC (JUST-N)
Inventor: KANTROWITZ M
Number of Countries: 090 Number of Patents: 002
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
WO 200033215	A1	20000608	WO 99US25686	A	19991101	200036	B
AU 200019073	A	20000619	AU 200019073	A	19991101	200044	

Priority Applications (No Type Date): US 98201569 A 19981130
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
WO 200033215	A1	E	16	G06F-017/30	

    Designated States (National): AE AL AM AT AU AZ BA BB BG BR BY CA CH CN
    CR CU CZ DE DK DM EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP
    KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX NO NZ PL PT RO RU SD SE
    SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW
    Designated States (Regional): AT BE CH CY DE DK EA ES FI FR GB GH GM GR
    IE IT KE LS LU MC MW NL OA PT SD SE SL SZ TZ UG ZW

| AU 200019073 | A | | | G06F-017/30 | Based on patent WO 200033215 |

Abstract (Basic): WO 200033215 A1
      NOVELTY - Occurrences of each **term** extracted from document is
   counted to establish a frequency value for each **term** . The characters
   in each **term** is counted. The frequency value for each **term** or
   monotonic function is multiplied by character count or monotonic
   function to form modulus for each **term** . The **terms** are **sorted**
   according to the moduli and moduli is accepted as **characteristic**
   keyword of the document's content.
      USE - In computer, world wide web for **term** **weighting** , for
   information **retrieval** applications such as document **retrieval** ,
   cross-language information **retrieval** , **keyword** extraction, document
   routing, classification, categorization, clustering, document
   filtering, **query** expansion, chapter, paragraph and sentence
   segmentation, spelling correction, **term** , **query** and document
   similarity metrics and text summarization.
      ADVANTAGE - Size of indexes in the information **retrieval**
   algorithm is reduced. Document summarized is easy to implement and use
   and requires only less memory. The method is scalable because it does
   not rely on information outside the document and so does not consume
   more resources as the number of documents increases. So the method is
   highly suitable for distributed information **retrieval** applications.
      DESCRIPTION OF DRAWING(S) - The figure shows the flow diagram
   explaining the computer program for implementing the characterizing
   terms extraction method.
         pp; 16 DwgNo 1/1
Title Terms: **TERM** ; EXTRACT; METHOD; COMPUTER; **SORT** ; EXTRACT; **TERM** ;
  ACCORD; GENERATE; MODULUS; ACCEPT; **TERM** ; GREATER; MODULUS;
  CHARACTERISTIC ; **KEYWORD** ; DOCUMENT; CONTENT
Derwent Class: T01
International Patent Class (Main): **G06F-017/30**
File Segment: EPI

07614748   **Image available**
INFORMATION   **RETRIEVING**   DEVICE,   INFORMATION   **RETRIEVING**   METHOD   AND
INFORMATION   **RETRIEVING**   PROGRAM

ABSTRACT

PROBLEM TO BE SOLVED: To solve a problem of requiring a long time for work
when the number of **retrieval** results increases for requiring a user to
successively check validity of the respective **retrieval** results since
order of these **retrieval** results does not always coincide with validity
in a **retrieval** purpose when obtaining a plurality of **retrieval** results.
SOLUTION: A character string having the same **attribute** as a **retrieving**
keyword and a character string having an **attribute** different from the
retrieving   **keyword**   are extracted   from a Web page gathered by a WWW
information gathering part 24, and a **score** of the Web page is set
according to a description degree of the character string to the Web page.

07001570   **Image available**
METHOD AND DEVICE FOR ELECTRONIC MAP **RETRIEVAL** AND RECORDING MEDIUM WITH
RECORDED ELECTRONIC MAP **RETRIEVING** PROGRAM

PUB. NO.:   2001-229182 [JP 2001229182 A]
PUBLISHED: August 24, 2001 (20010824)
INVENTOR(s):   FUKAYA KOJI
APPLICANT(s): HITACHI ENG CO LTD
APPL. NO.:   2000-039795 [JP 200039795]
FILED: February 14, 2000 (20000214)
INTL CLASS: **G06F-017/30** ; G06T-011/60; G09B-029/00; G09B-029/10;
               G01C-021/00

ABSTRACT

PROBLEM TO BE SOLVED: To provide an electronic map **retrieval** system which
has high **retrieval** precision and can correct a display position.

SOLUTION: This device is provided with a **retrieval** means 32 which
computes **property** data **retrieved** by decomposing a **retrieval** key **word**
 into elements and **weighting** them, and rates of matching, a candidate
plate list display means 33 which lists and displays the **retrieved**
 property data as candidate places in the decreasing **order** of the
matching rates, a map display means 34 which displays a map according to
the candidate place list, a position correcting means 35 which corrects the
position of a target item position mark displayed on the **retrieved** map,
and a correction position storage means 23 which stores the corrected
position on the map.

06807137   **Image available**
DOCUMENT CLASSIFICATION MANAGEMENT SYSTEM

ABSTRACT

PROBLEM TO BE SOLVED: To facilitate the document registration by automatic classification and to improve the hit rate at the time of document **retrieval** by comparing a user **word** in a document and the name of a classification node and classifying them into a corresponding classification node layers, and presenting the result to user.
SOLUTION: A classification node management module 101 which receives a document relative object generation request from a user interface module 100 generates a document relative object, sets the classification node ID of a relative source and the document ID of a relative destination as **properties** of the document relative object, and also sets a calculated or specified importance **score** . A document **retrieval** request is received from the module 100, the document relative object having the classification node ID of a **retrieval** object classification node is **retrieved** from a document management database and a document object having the document ID of the hit document relative object is **retrieved** from the document management database; and **retrieved** documents are **sorted** according to the importance **scores** and a **retrieval** hit document list is presented.

06097818   **Image available**
DOCUMENT **RETRIEVAL** DEVICE

PUB. NO.:   11-039337   [JP 11039337 A]
PUBLISHED: February 12, 1999 (19990212)
INVENTOR(s):   NOMOTO MASAKO
                 NOGUCHI NAOHIKO
                 SUGANO YUJI
                 SATO MITSUHIRO
                 INABA MITSUAKI
                 FUKUSHIGE TAKAO
APPLICANT(s):   MATSUSHITA ELECTRIC IND CO LTD
APPL. NO.:   09-198120   [JP 97198120]
FILED: July 24, 1997 (19970724)
INTL CLASS: **G06F-01**7/30

ABSTRACT

PROBLEM TO BE SOLVED: To show the effectiveness of a **word** and cooccurence
which are designated and to accurately **retrieve** a document which is
closer to a **retrieval** intention by extracting an **attribute** about an
appearance tendency as well as a **word** and cooccurence from a document.

SOLUTION: An input analyzing means 19 analyzes a **retrieval** condition,
segments a **word** and presents cooccurrence consisting of **words** that have
specific cooccurence relation to a user. A **word** collating means 21
collates each **word** that is extracted from the **retrieval** condition with
a **word** stored in a **word** frequency storing means 16 based on respective
weights of a **word** designated by the user and its appearance positional
level. A cooccurence information collating means 22 collates cooccurence
extracted from the **retrieval** condition with a cooccurence frequency
storing means 17 based on respective **weights** of cooccurence designated by
the user, its appearance positional level and its cooccurence level. A
document **order** deciding means 23 integrates collation results that are
performed by the means 21 and 22 in a document unit, decides the **order** of
each document and presents a result to the user through an input- output
controlling means.

05458664 **Image available**
SIMILAR INSTANCE **RETRIEVAL** DEVICE

PUB. NO.:   09-073464 [JP 9073464 A]
PUBLISHED: March 18, 1997 (19970318)
INVENTOR(s):   OKAMOTO AOSHI
                 SATO TAKESHI
APPLICANT(s):   FUJITSU LTD [000522] (A Japanese Company or Corporation), JP
                 (Japan)
APPL. NO.:   07-229774 [JP 95229774]
FILED: September 07, 1995 (19950907)
INTL CLASS: [6] **G06F-017/30**
JAPIO CLASS: 45.4 (INFORMATION PROCESSING -- Computer Applications)

ABSTRACT

PROBLEM TO BE SOLVED: To **retrieve** a similar instance in consideration of
the similarity of the **attribute** by registering **attribute** information in
a database, extracting a key **word** and taking the **attribute** information
out when a new instance is given, and outputting past instances in
decreasing **order** of the similarity.
SOLUTION: A **retrieval** system 1 extracts key **word** from the problem part
of a given instance to generate a key **word** number table 5, takes a key
word number out as to a key **word** extracted corresponding to an instance
number and registers it in a key **word** table 6, and registers the number
of key **word** numbers extracted from instances corresponding to category
numbers and the total numbers by categories in an instance quantity table
8, and an **attribute** information generating means 2 calculates **weight**
according to the instance quantity table 8. Then a similarity generating
means 3 generates similarity according to those pieces of information,
sorts the information in the decreasing **order** of the generated
similarity, and outputs the categories, similarity, etc., of past
instances.

03244676   **Image available**
ANALOGOUS KNOWLEDGE **RETRIEVING** METHOD USING GROUP KNOWLEDGE

ABSTRACT

PURPOSE: To newly obtain more accurate knowledges by replacing each partial
structure of a structure to be designed with another partial structure of a
group table in **order** to **retrieve** the analogous knowledges.
CONSTITUTION: A group table 4-5 is provided into an analogical processing
part 4 to accurately **retrieve** the analogous knowledges, and plural
partial structures have the same **characteristics** and same actions in
terms of the chemical reaction. Thus any partial structure has
approximately same chemical reactions within a group. The data on the group
are used for **retrieval** of the analogous knowledges. Thus it is regarded
that other partial structures included in the table 4-5 are analogous to
each other. Then the **score** of analogousness is improved and the effect of
analogy is also improved.

```
Set      Items    Description
S1     1964942    SEARCH?? OR SEARCHING OR FIND? ? OR FINDING OR FOUND OR RE-
                  TRIEVE? ? OR RETRIEVING OR RETRIEVAL OR QUERY OR QUERIES OR Q-
                  UERYING
S2      834262    KEYWORD? ? OR PHRASE? ? OR TERM? ? OR WORD? ?
S3      933758    ATTRIBUTE? ? OR CHARACTERISTIC? ? OR PROPERTY OR PROPERTIES
                  OR METADATA
S4      826121    SCALE? ? OR SCALING OR SCORE? ? OR SCORING OR WEIGHT?? OR -
                  WEIGHTING
S5     1352808    ORDER?? OR ORDERING OR SORT?? OR SORTING
S6        1320    S1 (30N) S2 (30N) S3 (30N) S4 (30N) S5
S8      261104    WEB OR WEBPAGE? ? OR WEBSITE? ? OR ONLINE OR ON()LINE OR I-
                  NTERNET? ? OR INTRANET? EXTRANET? ? OR WWW  OR WORLDWIDE()WEB
S9         113    S6 (30N) S8
S10         81    S9 AND IC=G06F
S11    1252008    FILE? ? OR DOCUMENT? ? OR ARTICLE? ? OR WEBPAGE? ? OR WEBS-
                  ITE?
S12      19899    S11 (3N) S3
S13         22    S1 (10N) S2 (30N) S12 (30N) S4 (30N) S5
S14         17    S13 AND IC=G06F
S15         17    IDPAT (sorted in duplicate/non-duplicate order)
S16         17    IDPAT (primary/non-duplicate records only)
File 348:EUROPEAN PATENTS 1978-2005/Dec W04
         (c) 2005 European Patent Office
File 349:PCT FULLTEXT 1979-2005/UB=20051229,UT=20051222
         (c) 2005 WIPO/Univentio
```

16/5,K/1   (Item 1 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2005 European Patent Office. All rts. reserv.

01826985
Document retrieval apparatus
Dokumentwiederauffingdungsvorrichtung
Appareil de recouvrement de documents
PATENT ASSIGNEE:
  KOKUSAI DENSHIN DENWA CO., LTD, (592872), 3-2, Nishi-shinjuku 2-Chome,
    Shinjuku-ku Tokyo, (JP), (Applicant designated States: all)
INVENTOR:
  Aoki, Keiko c/o KDDI R&D Laboratories Inc., 1-15 Ohara 2-chome,
    Kamifukuoka-shi Saitama, (JP)
  Matsumoto, Kazunori c/o KDDI R&D Laboratories Inc., 1-15 Ohara 2-chome,
    Kamifukuoka-shi Saitama, (JP)
  Hashimoto, Kazuo KDDI R&D Laboratories Inc., 1-15 Ohara 2-chome,
    Kamifukuoka-shi Saitama, (JP)
LEGAL REPRESENTATIVE:
  Skone James, Robert Edmund (50281), GILL JENNINGS & EVERY, Broadgate
    House, 7 Eldon Street, London EC2M 7LH, (GB)
PATENT (CC, No, Kind, Date): EP 1486891 A2   041215 (Basic)
                              EP 1486891 A3   050309
APPLICATION (CC, No, Date):   EP 2004022290 980211;
PRIORITY (CC, No, Date): JP 9741429 970212; JP 9767496 970306
DESIGNATED STATES: DE; FR; GB
RELATED PARENT NUMBER(S) - PN (AN):
  EP 859330 (EP 98301003)
INTERNATIONAL PATENT CLASS: **G06F-017/30**

ABSTRACT EP 1486891 A2
    A document retrieval apparatus is connected to the network, and
  comprises a cluster database (122) for storing a cluster of node
  information linked for clustering the documents to a hierarchical tree
  structure based on degree of similarity in all documents. The apparatus
  can post to the posted end address in the node information encountered on
  the way to follow links of the cluster by means of the cluster database
  when the document is updated. Also, the apparatus selects the specific
  number of documents, assigns non-selected documents respectively to a
  leaf node to be similar to the documents in the cluster, and indicates to
  repeat recursively the said operations toward a direction of the leaf
  node of cluster.
ABSTRACT WORD COUNT: 118
NOTE:
  Figure number on first page: 1

LEGAL STATUS (Type, Pub Date, Kind, Text):
 Application: 041215 A2 Published application without search report
 Examination: 041215 A2 Date of request for examination: 20041005
 Search Report: 050309 A3 Separate publication of the search report
 Change: 050420 A2 Inventor information changed: 20050228
 Change: 050504 A2 Inventor information changed: 20050317
LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200451	307
SPEC A	(English)	200451	4914
Total word count - document A			5221
Total word count - document B			0
Total word count - documents A + B			5221

INTERNATIONAL PATENT CLASS: **G06F-017/30**

...SPECIFICATION list, and pointers indicating parent and child nodes.

The frequency table of keywords lists by **weighting** with keywords based on the degree of similarity. The **order** of priorities is the descending of weighting points. The **weighting** points is the points counted by **weighting** the structure of the document and the occurrence frequency of **keywords** .

The frequency table is created as follows. First, the documents are cut down by limited **keywords** of a noun and an undefined **word** from entire text resource of a document by unit of morphological analysis. Then, the **keywords** are **weighted** . The **weighting** is reflected by not only the occurrence frequency of **keywords** , but also the tag structure of HTML (Hyper Text Makeup Language) text source. Thus, the frequency table showing a **characteristic** of the **document** can be provided.

The **weighting** with **keywords** in frequency table of the node information is sure to reflect the all documents positioned in a lower layer of the node. And the **retrieving** **keywords** are compared with the frequency tables of the child nodes, and a route passing through...

16/5,K/2   (Item 2 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2005 European Patent Office. All rts. reserv.

01784564
Image processing method and image processing system
Bildverarbeitungsverfahren und Bildverarbeitungssystem
Procede de traitement d'image et systeme de traitement d'image
PATENT ASSIGNEE:
  CANON KABUSHIKI KAISHA, (542366), 3-30-2 Shimomaruko, Ohta-ku, Tokyo,
    (JP), (Applicant designated States: all)
INVENTOR:
  Kaneda, Kitahiro, c/o Canon Kabushiki Kaisha 3-30-2, Shimomaruko,
    Ohta-ku, Tokyo, (JP)
  Tanioka, Hiroshi, c/o Canon Kabushiki Kaisha 3-30-2, Shimomaruko,
    Ohta-ku, Tokyo, (JP)
  Usami, Akihiro, c/o Canon Kabushiki Kaisha 3-30-2, Shimomaruko, Ohta-ku,
    Tokyo, (JP)
  Ohta, Ken-ichi, c/o Canon Kabushiki Kaisha 3-30-2, Shimomaruko, Ohta-ku,
    Tokyo, (JP)
  Ito, Hirohiko, c/o Canon Kabushiki Kaisha 3-30-2, Shimomaruko, Ohta-ku,
    Tokyo, (JP)
  Kato, Shinichi, c/o Canon Kabushiki Kaisha 3-30-2, Shimomaruko, Ohta-ku,
    Tokyo, (JP)
  Akiba, Tomohiro, c/o Canon Kabushiki Kaisha 3-30-2, Shimomaruko, Ohta-ku,
    Tokyo, (JP)
  Kanatsu, Tomotoshi, c/o Canon Kabushiki Kaisha 3-30-2, Shimomaruko,
    Ohta-ku, Tokyo, (JP)
  Misawa, Reiji, c/o Canon Kabushiki Kaisha 3-30-2, Shimomaruko, Ohta-ku,
    Tokyo, (JP)
  Terao, Yoshihide, c/o Canon Kabushiki Kaisha 3-30-2, Shimomaruko,
    Ohta-ku, Tokyo, (JP)
  Uzawa, Mitsuru, c/o Canon Kabushiki Kaisha 3-30-2, Shimomaruko, Ohta-ku,
    Tokyo, (JP)
LEGAL REPRESENTATIVE:
  Beresford, Keith Denis Lewis et al (28273), BERESFORD & Co. 16 High
    Holborn, London WC1V 6BX, (GB)
PATENT (CC, No, Kind, Date): EP 1455284 A2 040908 (Basic)
                              EP 1455284 A2 040908
                              EP 1455284 A3 050608
APPLICATION (CC, No, Date):   EP 2004250943 040220;
PRIORITY (CC, No, Date): JP 200344299 030221
DESIGNATED STATES: AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR;
  HU; IE; IT; LI; LU; MC; NL; PT; RO; SE; SI; SK; TR
EXTENDED DESIGNATED STATES: AL; HR; LT; LV; MK
INTERNATIONAL PATENT CLASS: G06K-009/20; **G06F-017/30**

ABSTRACT EP 1455284 A2
    This invention provides an image processing method which allows easy
  re-use of image information that is stored to minimize deterioration of
  image quality and the storage capacity. Storage means is searched for
  original digital data corresponding to each input image. If no original
  digital data is found, the input image is converted into vector data, and
  is stored as digital data in the storage means. A sheet including at
  least one of information associated with the found original digital data
  when the original digital data is found in the search step and
  information associated with digital data which is obtained by converting
  the image into the vector data in the vectorization step and is stored in
  the storage step when no original digital data is found in the search
  step is generated, thus providing a sheet that allows easy re-use.
ABSTRACT WORD COUNT: 141
NOTE:
  Figure number on first page: 19

LEGAL STATUS (Type, Pub Date, Kind, Text):
  Application: 040908 A2 Published application without search report

...INTERNATIONAL PATENT CLASS: **G06F-017/30**

...SPECIFICATION layout information and font information extracted by the
aforementioned method.
  As a method of extracting **keywords** from text data contained in a
document, for example, the entire text data is decomposed into **words**
by, among others, morphological analysis. All **words** are **sorted** in
accordance with their frequencies of use, and are selected as **keywords**
in descending **order** of frequency of use. In **order** to extract more
effective **keywords** , **words** may be compared with a **keyword** database,
which is prepared in advance.
  As for information of an ID, date, and author, if a file is **found** by
a digital file **search** process, such information is acquired as
property information of that **file** .
  As for abstract information, the following method of generating an
abstract of text data formed...

...method of calculating the importance level, a method of calculating the
frequencies of occurrence of **words** contained in the entire text data,
giving a high score to a **word** that appears frequently, and calculating
the importance level of each sentence or clause as a...

...and font information to increase the importance level of that sentence,
or to increase the **scores** of words included in that sentence, and the
like may be used. Finally, an abstract...

16/5,K/3   (Item 3 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2005 European Patent Office. All rts. reserv.

01697975
Document retrieval system and question answering system
Dokumentwiederauffindungssystem und Frage-Antwortsystem
Systeme de recouvrement de documents et systeme de questions/reponses
PATENT ASSIGNEE:
  MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD., (216883), 1006, Oaza-Kadoma,
    Kadoma-shi, Osaka 571-8501, (JP), (Applicant designated States: all)
INVENTOR:
  Nomoto, Masako, 7-18-504, Minamisumiyoshi, Tokorozawa-shi, Saitama
    359-1125, (JP)
  Sato, Mitsuhiro, 4-11-11, Miyanosato, Atsugi-shi, Kanagawa 243-0216, (JP)
  Suzuki, Hiroyuki, 4-22-10-501, Shimokodanaka, Nakahara-ku, Kawasaki-shi,
    Kanagawa 211-0041, (JP)
LEGAL REPRESENTATIVE:
  Grunecker, Kinkeldey, Stockmair & Schwanhausser Anwaltssozietat (100721)
  , Maximilianstrasse 58, 80538 Munchen, (DE)
PATENT (CC, No, Kind, Date): EP 1391834 A2 040225 (Basic)
APPLICATION (CC, No, Date):   EP 2003018569 030818;
PRIORITY (CC, No, Date): JP 2002238031 020819; JP 2003189111 030630
DESIGNATED STATES: AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR;
  HU; IE; IT; LI; LU; MC; NL; PT; RO; SE; SI; SK; TR
EXTENDED DESIGNATED STATES: AL; LT; LV; MK
INTERNATIONAL PATENT CLASS: **G06F-017/30**

ABSTRACT EP 1391834 A2
    A document retrieval system capable of obtaining information requested
  by the user with a high degree of accuracy. In this system, the query
  input section 102 receives query input by the user. The keyword
  extraction section 104 analyzes the input query and extracts keywords.
  The keyword type assignment section 106 decides the type of each
  extracted keyword and assigns a keyword type. The question type decision
  section 108 decides the question type. The keyword classification section
  110 classifies the keywords to which the keyword types are assigned into
  a major type and minor type with reference to the keyword classification
  rules stored in the keyword classification rule storage section 112. The
  document retrieval section 114 searches a document collection stored in
  the document storage section 116 using the classified keyword groups and
  obtains the document of the retrieved result.
ABSTRACT WORD COUNT: 140
NOTE:
  Figure number on first page: 1

LEGAL STATUS (Type, Pub Date, Kind, Text):
 Application: 040225 A2 Published application without search report
LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200409	2550
SPEC A	(English)	200409	10113
Total word count - document A			12663
Total word count - document B			0
Total word count - documents A + B			12663

INTERNATIONAL PATENT CLASS: **G06F-017/30**

...SPECIFICATION only documents including a tag <LOCATION> are
  extracted.
    At this time, when the **search** question type as a result of the
  decision by the question type decision section 108...

...into account minor keywords, too, while limiting the search range to only documents including major **keywords** and having semantic attributes that match the search question type and obtain a retrieved result...

...has described the case where a search method using the search question type and semantic **attributes** in **documents** is combined with the first search method in Embodiment 1 shown in FIG.5 as...

...shown in FIG.6 or the third search method (ranking by layer including restrictiveness of **keywords**) in Embodiment 1 shown in FIG.8.
Furthermore, this embodiment carries out a search in...

...this embodiment has described the case where the semantic attribute assignment section 202 assigns semantic **attributes** to **document** collections beforehand, as an example, but this embodiment is not limited to this and can also be adapted so as to assign semantic **attributes** to only **document** collections obtained after searching for document collections. It generally takes a considerable calculation time to...

...number of documents, and therefore adopting such a configuration makes it possible to assign semantic **attributes** to only necessary **documents** and streamline the processing.
Furthermore, this embodiment can also be adapted so as to search for **documents** whose semantic **attribute** values are normalized (**document** collection with normalized semantic **attributes**) as **document** collections. In this case, when, for example, "2000/6/30" is specified as a keyword documents to be **retrieved** to a

01321702
System and method for information processing
System und Verfahren zur Informationsverarbeitung
Systeme et procede pour le traitement d'informations
PATENT ASSIGNEE:
  SONY CORPORATION, (214021), 7-35 Kitashinagawa 6-chome Shinagawa-ku,
    Tokyo 141, (JP), (Applicant designated States: all)
INVENTOR:
  Saito, Mari, c/o Sony Corporation, 7-35, Kitashinagawa 6-chome,
    Shinagawa-ku, Tokyo, (JP)
  Yamamoto, Noriyuki, c/o Sony Corporation, 7-35, Kitashinagawa 6-chome,
    Shinagawa-ku, Tokyo, (JP)
  Hourin, Hiroyuki, c/o Sony Corporation, 7-35, Kitashinagawa 6-chome,
    Shinagawa-ku, Tokyo, (JP)
  Ohmura, Kazunori, c/o Sony Corporation, 7-35, Kitashinagawa 6-chome,
    Shinagawa-ku, Tokyo, (JP)
LEGAL REPRESENTATIVE:
  Korber, Martin, Dipl.-Phys. et al (88321), Mitscherlich & Partner
    Patentanwalte Sonnenstrasse 33, 80331 Munchen, (DE)
PATENT (CC, No, Kind, Date): EP 1128276 A1   010829 (Basic)
APPLICATION (CC, No, Date):   EP 2001104031 010220;
PRIORITY (CC, No, Date): JP 200042303 000221; JP 200042305 000221; JP
    2000187152 000622
DESIGNATED STATES: DE; FR; GB
EXTENDED DESIGNATED STATES: AL; LT; LV; MK; RO; SI
INTERNATIONAL PATENT CLASS: **G06F-017/27 ;   G06F-017/30**

ABSTRACT EP 1128276 A1
    Disclosed is an information processing apparatus, an information
processing method, and a program storage medium which can present
associated information related to a document to be processed to a user.
An accumulation block accumulates a database of associated information. A
presentation block presents to the user the associated information
corresponding to the document to be processed at occurrence of an event.
An agent control block controls the manner of displaying an agent for
example.
ABSTRACT WORD COUNT: 76
NOTE:
  Figure number on first page: 1

LEGAL STATUS (Type, Pub Date, Kind, Text):
 Application: 010829 A1 Published application with search report
 Examination: 020403 A1 Date of request for examination: 20020124
 Examination: 041222 A1 Date of dispatch of the first examination
                           report: 20041108
LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200135	1232
SPEC A	(English)	200135	9553
Total word count - document A			10785
Total word count - document B			0
Total word count - documents A + B			10785

INTERNATIONAL PATENT CLASS: **G06F-017/27 ...**

... G06F-017/30

...SPECIFICATION processing block 3 and performs a morphemic analysis on
  the extracted text data to extract **keywords** . The document content

processing block 4 obtains the occurrence frequency of the **keywords** and
the distribution status over plural documents and computes the weight of
the **keyword** of each document group by use of the tf (center dot) idf
method for example...

...preparation block 5 creates a database of the attribute information and
the weights of all **keywords** included in each of the documents grouped
by the **document attribute** processing block 3. To be more specific, as
shown in FIG. 4, the grouped documents are sorted in a time dependent
manner and then the weights of all **keywords** 1 through n included in the
grouped documents are sorted in a time dependent manner...

...in the storage block 29. In FIG. 4, weight A1 denotes the weight value
of **keyword** 1 in document A and weight B2 denotes the weight value of
keyword 2 in document B for example. Further, if **keyword** 1 is not
included in document B, weight B1 becomes 0.
    In step S6, the...

...5 selects a keyword with its weight being higher than a predetermined
threshold as a **search** keyword (an important word) and selects the
number of keywords specified in the descending **order** of weights,
supplying the selected keywords to the associated information **retrieval**
block 6. By use of the **search** keyword supplied from the document
feature database preparation block 5 as a **search** condition, the
associated information **retrieval** block 6 accesses a **search** engine on
the Internet to retrieve search results and outputs the URL and title of

01289039
**Document search system with automatic field selection and field controlled
    document ranking**
**Dokumentweiderauffindunssystem   mit   automatischen   Feldauswahl   und
    feldgesteuerte Dokumentordnung**
**Systeme de recherche de documnets avec selection de zones automatique et
    rangement de documents controllee par zones**
PATENT ASSIGNEE:
  Eidetica B.V., (3197070), Twinning Center (Matrix II), Kruislaan 400,
    1098 SM Amsterdam, (NL), (Applicant designated States: all)
INVENTOR:
  Groenink, Annius Victor, 684, Keizersgracht, 1017 ET Amsterdam, (NL)
LEGAL REPRESENTATIVE:
  Brookhuis, Hendrik Jan Arnold (79631), van Exter Polak & Charlouis B.V.
    P.O. Box 3241, 2280 GE Rijswijk, (NL)
PATENT (CC, No, Kind, Date): EP 1107133 A1   010613 (Basic)
APPLICATION (CC, No, Date):   EP 2000204386 001207;
PRIORITY (CC, No, Date): NL 1013793 991208
DESIGNATED STATES: DE; GB; NL
EXTENDED DESIGNATED STATES: AL; LT; LV; MK; RO; SI
INTERNATIONAL PATENT CLASS: **G06F-017/30**

ABSTRACT EP 1107133 A1
    A document search system comprising data storage means, that contains
at least a metadata collection having a collection of three-tuples
(<metadata>, <fieldtype-id>, <document-idlist>), which
metadata collection is obtained from a collection of documents and is
composed of a number of fields, whereby a fieldtype-identifier
<fieldtype-id> is assigned to each field, and in which metadata
collection each three-tuple indicates that for all documents in the
non-empty list of document identifiers <doc-id-list> the element
<metadata> is metadata of a field identified by
<fieldtype-id>.
    Further the search system comprises a search algorithm with a matching
algorithm having as input a query, which query comprises at least an
enumeration of pairs ((<target>, <weight>),
<fieldtype-idlist>), in which pairs <weight> is a real number
on the interval (0;1), and which matching algorithm has the metadata
collection as input, and compares per <fieldtype-id> the values of
<metadata> to the values of <target> in the query and
includes <weight> in the comparison, and which matching algorithm
has as output a relevance collection comprising three-tuples
(<target>, <fieldtype-id>, <doc-idlist>), which
relevance collection contains per unique combination of <metadata>
and <fieldtype-id> a list of document identifiers
<doc-idlist> in which the identifiers identify documents that are
considered sufficiently relevant with respect to the query by the
matching algorithm.
ABSTRACT WORD COUNT: 207
NOTE:
  Figure number on first page: 1

LEGAL STATUS (Type, Pub Date, Kind, Text):
 Application: 010613 A1 Published application with search report
 Examination: 020417 A1 Date of request for examination: 20011009
 Assignee: 051228 A1 Transfer of rights to new applicant: Stork
                           Aerospace Group B.V. (7187650) 4, Industrieweg
                           3351 LB Papendrecht NL
LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:

```
Available Text  Language    Update     Word Count
     CLAIMS A   (English)   200124        776
     SPEC A     (English)   200124       3571
Total word count - document A            4347
Total word count - document B               0
Total word count - documents A + B       4347
```

INTERNATIONAL PATENT CLASS: **G06F-017/30**

...SPECIFICATION the input means, without a change of one of the elements
(<target>,< **weight** >). This has as advantage that the user is
enabled to make the search algorithm...

...the field types present in the metadata collection for each combination
(<target>,< **weight** >) in <field type id-list>, in
response to an addition via the input means of at least one target in the
query . After an addition of a target, the user will usually **search**
again in all field types, and only after that he will want to select one
...

...the output means all the document identifiers that are present in the
filtered relevance collection, **sorted** by criteria based on the data
given by the relevance collection on the individual field **metadata** . As
such, the **documents** having the highest predicted relevance can be
arranged on top of the list, and therefore easily be distinguished from
the other documents.
   The **sorting** thereby is advantageously done according to one of the
function values r1, r2, r3, and...

...claim 9 it is noted that the field length may be calculated differently
in each **search** system. If the notion of **term** is defined as a single
word , the field length equals the number of **words** in the field. If the
notion of **term** is defined as the words indicating a concept, the field
length equals the number of...

...twice, viz. once as algebra and once as linear algebra.
   It is particularly advantageous to **sort** not according to a single
criterion as mentioned above, but according to a staged **sorting**
algorithm in which the number of stages is at least two and in which one

16/5,K/8   (Item 8 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2005 European Patent Office. All rts. reserv.

00947222
Document retrieval apparatus
Dokumentwiederauffindungsvorrichtung
Appareil de recouvrement de documents
PATENT ASSIGNEE:
  KOKUSAI DENSHIN DENWA CO., LTD, (592872), 3-2, Nishi-shinjuku 2-Chome,
    Shinjuku-ku Tokyo, (JP), (applicant designated states: DE;FR;GB)
INVENTOR:
  Aoki, Keiko, c/o KDDI R&D Laboratories Inc. 1-15 Ohara 2-chome,
    Kamifukuoka-shi Saitama, (JP)
  Matsumoto, Kazunori, c/o KDDI R&D Laboratories Inc. 1-15 Ohara 2-chome,
    Kamifukuoka-shi Saitama, (JP)
  Hashimoto, Kazuo c/o KDDI R&D Laboratories Inc., 1-15 Ohara 2-chome,
    Kamifukuoka-shi Saitama, (JP)
LEGAL REPRESENTATIVE:
  Skone James, Robert Edmund (50281), GILL JENNINGS & EVERY Broadgate House
    7 Eldon Street, London EC2M 7LH, (GB)
PATENT (CC, No, Kind, Date): EP 859330 A1 980819 (Basic)
APPLICATION (CC, No, Date):   EP 98301003 980211;
PRIORITY (CC, No, Date): JP 9741429 970212; JP 9767496 970306
DESIGNATED STATES: DE; FR; GB
RELATED DIVISIONAL NUMBER(S) - PN (AN):
    (EP 2004022290)
INTERNATIONAL PATENT CLASS: **G06F-017/30**

ABSTRACT EP 859330 A1
  A document retrieval apparatus is connected to the network, and
  comprises a cluster database (122) for storing a cluster of node
  information linked for clustering the documents to a hierarchical tree
  structure based on degree of similarity in all documents. The apparatus
  can post to the posted end address in the node information encountered on
  the way to follow links of the cluster by means of the cluster database
  when the document is updated. Also, the apparatus selects the specific
  number of documents, assigns non-selected documents respectively to a
  leaf node to be similar to the documents in the cluster, and indicates to
  repeat recursively the said operations toward a direction of the leaf
  node of cluster.
ABSTRACT WORD COUNT: 118

LEGAL STATUS (Type, Pub Date, Kind, Text):
Examination:	040630	A1	Date of dispatch of the first examination report: 20040514
Application:	980819	A1	Published application (A1with Search Report ;A2without Search Report)
Change:	050504	A1	Inventor information changed: 20050316
Change:	041110	A1	Application number of divisional application (Article 76) changed: 20040922
Change:	050420	A1	Inventor information changed: 20050228
Examination:	980819	A1	Date of filing of request for examination: 980227
Change:	990506	A1	Designated Contracting States (change)

LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:
Available Text	Language	Update	Word Count
CLAIMS A	(English)	9834	1141
SPEC A	(English)	9834	5028
Total word count - document A			6169
Total word count - document B			0
Total word count - documents A + B			6169

...SPECIFICATION list, and pointers indicating parent and child nodes.
   The frequency table of keywords lists by   **weighting**   with keywords
based on the degree of similarity. The   **order**   of priorities is the
descending of weighting points. The   **weighting**   points is the points
counted by   **weighting**   the structure of the document and the occurrence
frequency of   **keywords** .
   The frequency table is   created as follows. First, the documents are
cut down by limited   **keywords**   of a noun and an undefined   **word**   from
entire text resource of a document by unit of morphological analysis.
Then, the   **keywords**   are   **weighted** . The   **weighting**   is reflected by not
only the occurrence frequency of   **keywords** , but also the tag structure
of HTML (Hyper Text Makeup Language) text source. Thus, the frequency
table showing a   **characteristic**   of the   **document**   can be provided.
   The   **weighting**   with   **keywords**   in frequency table of the node
information is sure to reflect the all documents positioned in a lower
layer of the node. And the   **retrieving**   **keywords**   are compared with the
frequency tables of the child nodes, and a route passing through...

16/5,K/9   (Item 9 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2005 European Patent Office. All rts. reserv.

00433723
A method of information retrieval for a database system
Verfahren zur Informations-Wiedergewinnung fur ein Datenbanksystem
Methode de recherche documentaire pour un systeme de base de donnees
PATENT ASSIGNEE:
   International Business Machines Corporation, (200120), Old Orchard Road,
      Armonk, N.Y. 10504, (US), (applicant designated states: DE;FR;GB)
INVENTOR:
   Mitsui, Kinichi, 148 Futako, Takatsu-ku, Kawasaki-shi, Kanagawa-ken, (JP)
LEGAL REPRESENTATIVE:
   Moss, Robert Douglas (34141), IBM United Kingdom Limited Intellectual
      Property Department Hursley Park, Winchester Hampshire SO21 2JN, (GB)
PATENT (CC, No, Kind, Date): EP 420424   A2   910403 (Basic)
                              EP 420424   A3   921202
                              EP 420424   B1   971203
APPLICATION (CC, No, Date):   EP 90309707 900905;
PRIORITY (CC, No, Date): JP 89242421 890920
DESIGNATED STATES: DE; FR; GB
INTERNATIONAL PATENT CLASS: **G06F-017/30**
CITED PATENTS (EP A): DE 3901485 A; US 4358824 A
CITED REFERENCES (EP A):
   JOURNAL OF CHEMICAL INFORMATION AND COMPUTER SCIENCES vol. 15, no. 1,
      February 1975, pages 32 - 39 H. S. HEAPS 'Data Compression of Large
      Document Data Bases';

ABSTRACT EP 420424 A2
   A database comprises a sequential file 15 and a transposed file 11 in
external memory. A query specifies various keywords, which are used to
access the transposed file. Data concerning the various keywords, and
especially identifiers for retrieval objects containing the keywords, are
loaded into main storage. Inside main storage this information is
rearranged, so as to be able to rank the retrieval objects in terms the
keywords they contain. Identifiers from this ordered list can then be
used to access information from the sequential file, with only retrieval
objects with high query scores being accessed.   (see image in original
document)
ABSTRACT WORD COUNT: 105

LEGAL STATUS (Type, Pub Date, Kind, Text):
 Application: 910403 A2 Published application (A1with Search Report
                              ;A2without Search Report)
 Examination: 910403 A2 Date of filing of request for examination:
                              901213
 Search Report: 921202 A3 Separate publication of the European or
                              International search report
 Examination: 960626 A2 Date of despatch of first examination report:
                              960514
 Grant: 971203 B1 Granted patent
 Lapse: 981028 B1 Date of lapse of the European patent in a
                              Contracting State: FR 980430
 Oppn None: 981125 B1 No opposition filed
LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS B	(English)	9711W4	679
CLAIMS B	(German)	9711W4	677
CLAIMS B	(French)	9711W4	775
SPEC B	(English)	9711W4	5344
Total word count - document A			0
Total word count - document B			7475
Total word count - documents A + B			7475

...SPECIFICATION numerical attributes
    Next, a method for utilizing a numerical attribute like publication
year for quantitative **retrieval** is described, with reference to Figure
5. When calculating the score relative to a numerical...

...value to a certain value, the higher is its score.
    With regard to the numerical **attribute** , a transposed **file** which
relates each **attribute** value to each **retrieval** object may be prepared
similarly as for **keywords** . In Figure 5, K3 denotes a numerical value
attribute representing publication year (41). On the transposed file, the
retrieval objects are arranged in the ascending or descending **order**
of the attribute values, permitting high speed access in the ascending or
descending **order** . This may be attained by making use of an existing
technique like B-blocks.
    Now, when a numerical attribute is generally used as a **retrieval**
condition, the range of values influencing the **score** is wide, so that
it is unavoidable to frequently access the external storage to access the
overall range of the values influencing the **score** on the transposed
file. When outputting only objects with high **scores** , there is a
possibility that it is sufficient to access only the parts which give
high **scores** where the record of the numerical attribute is concerned.
    With regard to the access key...
...the numerical attribute, by utilizing a transposed file on which the
retrieval object identifiers are **sorted** in the attribute **order**
beforehand, the external storage is accessed sequentially from the part
where the highest **score** may be obtained and, then, the retrieval may be
ended at a point when the..

16/5,K/10   **(Item 10 from file: 349)**
DIALOG(R)File 349:PCT FULLTEXT
(c) 2005 WIPO/Univentio. All rts. reserv.

01304210   **Image available**
SYSTEM AND METHOD FRO DYNAMICALLY GENERATING A SELECTABLE SEARCH EXTENSION
SYSTEME ET METHODE POUR GENERER DYNAMIQUEMENT UNE EXTENSION DE RECHERCHE
     SELECTIONNABLE
Patent Applicant/Assignee:
  MICROSOFT CORPORATION, One Microsoft Way, Redmont, WA 98052, US, US
    (Residence), US (Nationality)
Inventor(s):
  ANTHONY Colin R, 11207 123rd Lane, NE Apt. C-21, Kirkland, WA 98033, US,
  VAN DOK Cornelis K, 16818 S.E. 2nd Place, Bellevue, WA 98008, US,
  IWEMA Marieke, 10300 Rainer Avenue S., Seattle, WA 98178, US,
Legal Representative:
  STROHM Scott B (et al) (agent), Shook, Hardy & Bacon L.L.P., 2555 Grand
    Blv., Kansas City, MI 64108-2613, US,
Patent and Priority Information (Country, Number, Date):
  Patent: WO 2005111868 A2 20051124 (WO 05111868)
  Application: WO 2004US24634 20040730 (PCT/WO US04024634)
  Priority Application: US 2004566947 20040503
Designated States:
(All protection types applied unless otherwise stated - for applications
2004+)
  AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM
  DZ EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC
  LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NA NI NO NZ OM PG PH PL PT RO
  RU SC SD SE SG SK SL SY TJ TM TN TR TT TZ UA UG UZ VC VN YU ZA ZM ZW
  (EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PL PT RO
  SE SI SK TR
  (OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
  (AP) BW GH GM KE LS MW MZ NA SD SL SZ TZ UG ZM ZW
  (EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: **G06F-017/30**
Publication Language: English
Filing Language: English
Fulltext Availability:
  Detailed Description
  Claims
Fulltext Word Count: 4293

English Abstract
  A system and related techniques accept user-inputted search terms, for
  example to perform a search for files or other data or objects.
  Corresponding matches to those terms may be presented to the user in a
  "word-wheel"-type breakout list generated on the fly for groupings of
  hits by attributes or other criteria, as the system searches through the
  file system at the current level or point in the file system hierarchy.
  According to embodiments, when the search logic fails to locate a hit on
  the inputted search term at the current level or point in the file system
  hierarchy, an extension of the search to different levels or points in
  the file system hierarchy may be automatically generated, and for
  instance presented to the user as a selectable search box. That box may
  for example be highlighted to the user for easy selection. When the user
  does select the selectable search box, the user's search, for instance
  for files of type or extension ".doc" or ".memo", may be seamlessly
  extended to other files, folders, trees or other points or levels in the
  file system hierarchy. Search results may be continuously or dynamically
  updated as the user, for example, enters more characters or other data.

French Abstract
  L'invention concerne un systeme et des techniques associees acceptant des
  termes de recherche saisis par un utilisateur pour effectuer, par
  exemple, une recherche de fichiers ou d'autres donnees ou d'autres

objets. Les objets correspondant a ces termes peuvent etre presentes a
l'utilisateur dans une liste thematique de type "word wheel", generee a
la volee, pour des groupements de correspondances, par attributs ou par
autres criteres, alors que le systeme effectue des recherches dans le
systeme de fichiers, au niveau ou au point actuel de la hierarchie de
systemes de fichiers. En fonction des modes de realisation de
l'invention, lorsque la logique de recherche n'arrive pas a localiser une
correspondance pour le terme de recherche saisi sur le niveau actuel ou
sur le point actuel de la hierarchie de systemes de fichiers, une
extension de la recherche sur des niveaux ou des points differents de la
hierarchie de systemes de fichiers peut etre automatiquement generee, et
par exemple, peut etre presentee a l'utilisateur, en tant que pave de
recherche selectionnable. Ce pave peut, par exemple, etre mis en evidence
pour l'utilisateur, pour faciliter sa selection. Lorsque l'utilisateur
selectionne le pave de recherche selectionnable, la recherche de
l'utilisateur, par exemple de fichiers de type ou d'extension ".doc" ou
".memo", peut etre etendue sans coupure a d'autres fichiers, dossiers,
arbres ou a d'autres points ou a d'autres niveaux de la hierarchie de
systemes de fichiers. Les resultats de recherche peuvent etre mis a jour
de maniere continue ou dynamique, lorsque l'utilisateur saisit, par
exemple, plus de caracteres ou encore d'autres donnees.

Legal Status (Type, Date, Text)
Publication 20051124 A2 Without international search report and to be
                         republished upon receipt of that report.
Main International Patent Class: **G06F-017/30**
Fulltext Availability:
  Detailed Description

Detailed Description
... file size, or date created or modified. A user may at times also
   choose to **search** for files based on internal file'corii@nt','such as
   desired text or numbers. The...

...within a large corporation or other organization. In other cases, a user
   may wish to **sort** or **search** through a collection or catalogue of
   musical, video or other media or file material. Some **search** tools and
   facilitieš have evolved in response to large- **scale** file **search** and
   other requirements.

   For example, some applications and other packages may present
   the user with an input box type of **search** interface, where the user may
   enter **search** **terms** such as **file** extensions or other **attributes** ,
   or in- **file** characters or text. As the **search** , for example through a
   local hard drive and associated file system, progresses, files which
   partly...
...attiributes or text may be displayed to the user to select or
   manipulate.

   However, existing **search** tools may be constrained by certain
   limitations in usability or functionality. For instance, even such
   search tools as exist merely present the results gathered from
   searching the client or other file system at the current level or point
   in the file system hierarchy. So if no results are **found** in a given
   directory or folder, the user may be required to restart and reenter...

DIALOG(R)File 349:PCT FULLTEXT

01240860   **Image available**
SEARCH ENGINE WITH HIERARCHICALLY STORED INDICES
MOTEUR DE RECHERCHE AVEC INDEX A MEMORISATION HIERARCHIQUE
Patent Applicant/Assignee:
  OVERTURE SERVICES INC, 74 North Pasadena Avenue, Pasadena, CA 91103, US,
    US (Residence), US (Nationality), (For all designated states except:
    US)
Patent Applicant/Inventor:
  RISVIK Knut Magne, Sigrid Undsets vei 27, N-7023 Trondheim, NO, NO
    (Residence), NO (Nationality), (Designated only for: US)
  AASHEIM Yngve, Nedre Mollenberggate 76A, N-7043 Trondheim, NO, NO
    (Residence), NO (Nationality), (Designated only for: US)
  EGGE Tor, Fjordgata 16A, N-7010 Trondheim, NO, NO (Residence), NO
    (Nationality), (Designated only for: US)
  PETTERSEN Havard, Elgesetergate 26A, N-7030 Trondheim, NO, NO (Residence)
    , NO (Nationality), (Designated only for: US)
Legal Representative:
  RUBIN Steven (agent), Brown Raysman Millstein Felder & Steiner LLP, 900
    Third Avenue, New York, NY 10022, US,
Patent and Priority Information (Country, Number, Date):
  Patent: WO 200548069 A2 20050526 (WO 0548069)
  Application: WO 2004US37507 20041109 (PCT/WO US04037507)
  Priority Application: US 2003705641 20031110
Designated States:
(All protection types applied unless otherwise stated - for applications
2004+)
  AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM
  DZ EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC
  LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NA NI NO NZ OM PG PH PL PT RO
  RU SC SD SE SG SK SL SY TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW
  (EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LU MC NL PL PT
  RO SE SI SK TR
  (OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
  (AP) BW GH GM KE LS MW MZ NA SD SL SZ TZ UG ZM ZW
  (EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: **G06F**
Publication Language: English
Filing Language: English
Fulltext Availability:
  Detailed Description
  Claims
Fulltext Word Count: 5101
English Abstract
  A search engine comprising a crawler which crawls the WWW and stores
  pages found on the WWW in a database. An indexer indexes the pages in the
  database to produce a primary index. A document mapping section maps
  pages in the database into a plurality of tiers based on a ranking of the
  pages. The ranking may be based on portions of the pages which have a
  relatively higher value context. A processor produces a plurality of
  sub-indices from the primary index based on the mapping. The sub-indices
  are stored in a search node cluster. The cluster is a matrix of search
  nodes logically arranged in a plurality of rows and columns. Search nodes
  in the same column include the same sub-index. Search nodes in the same
  row include distinct sub-indices. A search query received by a user is
  sent to a dispatcher which, in turn, forwards the query to the first tier
  of search nodes. A fall through algorithm is disclosed which indicates
  when the dispatcher should forward the search query to other tiers of
  search nodes.

French Abstract
  Cette invention se rapporte a un moteur de recherche qui comprend un

moteur de recherche Web qui parcourt le Web et memorise les pages
trouvees sur le Web dans une base de donnees. Un indexeur indexe les
pages dans la base de donnees pour produire un index primaire. Une
section de cartographie de documents cartographie les pages dans la base
de donnees en plusieurs niveaux sur la base d'un classement des pages. Ce
classement peut etre base sur les parties des pages qui presentent un
contexte de valeur relativement superieure. Un processeur produit
plusieurs sous-index a partir de l'index primaire sur la base de la
cartographie. Les sous-index sont memorises dans une grappe de noeuds de
recherche. Cette grappe est constituee par une matrice de noeuds de
recherche agences logiquement en plusieurs rangees et colonnes. Les
noeuds de recherche de la meme colonne ont le meme sous-index. Les noeuds
de recherche de la meme rangee ont des sous-index distincts. Une
interrogation de recherche recue par un utilisateur est envoyee a un
repartiteur qui, a son tour, transmet l'interrogation au premier niveau
des noeuds de recherche. Cette invention contient un algorithme a
transfert implicite qui indique a quel moment le repartiteur doit
transferer l'interrogation de recherche a d'autres niveaux des noeuds de
recherche.

Legal Status (Type, Date, Text)
Publication 20050526 A2 Without international search report and to be
                         republished upon receipt of that report.

Main International Patent Class: **G06F**
Fulltext Availability:
  Detailed Description

Detailed Description
... a plurality of sub-indices (discussed below) and each sub-index is
    sent to a **search** node in a search node cluster 106.

    [00041 In use, a user I 1 2...

...cluster 106 search respective parts of the primary index produced by
    indexer 104 and return **sorted** search results along with a document
    identifier and a **score** to dispatcher I IO. Dispatcher 1 1 0 merges the
    received results to produce a final list displayed to the users 1 12
    sorted by relevance **scores** . The relevance **score** is a function of the
    query itself and the type of document produced. Factors that are used for
    relevance include: a static relevance **score** for the document such as
    link cardinality and page quality, superior parts of the **document** such
    as titles, **metadata** and **document** headers, authority of the document
    such as external references and the "level" of the references, and
    document statistics such as query **term** frequency in the document,
    global **term** frequency, and **term** distances within the document.
    [00051 Referring now to Fig. 2, a cluster 106 of search...

...column 122 of search nodes, the same set of indices is replicated for
    each respective **search** node. For example, the **search** node in column
    122a, row 124a, includes the
    2
    same subset of indices as the **search** node in column 122a, 124b. In each
    row 124 of **search** nodes, a different subset of indices is used. The
    indices are equally split so as...

01191967   **Image available**
DATABASE QUERY USER INTERFACE
INTERFACE-UTILISATEUR D'INTERROGATION DE BASE DE DONNEES
Patent Applicant/Assignee:
  MICROSOFT CORPORATION, One Microsoft Way, Redmond, WA 98052, US, US
    (Residence), US (Nationality), (For all designated states except: US)
Inventor(s):
  VRONAY David P, 3 F, Beijing Sigma Center, No. 49, Xhichun Road, Haidian
    District, Beijing100080, CN,
  MARCJAN Cezary, 15916 NE 40th Way, Redmond, WA 98052, US,
  TURSKI Andrzej, 5708 223rd Avenue, NE, Redmond, WA 98053, US,
  KOTT Ryszard, 16717 NE 98th Pl., Redmond, WA 98052, US,
Legal Representative:
  AMIN Himanshu S (et al) (agent), Amin & Turocy, LLP, 1900 E. 9th Street,
    24th Floor, National City Center, Cleveland, OH 44115, US,
Patent and Priority Information (Country, Number, Date):
  Patent: WO 2004114062 A2 20041229 (WO 04114062)
  Application: WO 2004US18503 20040610 (PCT/WO US04018503)
  Priority Application: US 2003461832 20030613
Designated States:
(All protection types applied unless otherwise stated - for applications
2004+)
  AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM
  DZ EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC
  LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NA NI NO NZ OM PG PH PL PT RO
  RU SC SD SE SG SK SL SY TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW
  (EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PL PT RO
  SE SI SK TR
  (OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
  (AP) BW GH GM KE LS MW MZ NA SD SL SZ TZ UG ZM ZW
  (EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: **G06F**
Publication Language: English
Filing Language: English
Fulltext Availability:
  Detailed Description
  Claims
Fulltext Word Count: 8409

English Abstract
  A database query user interface combines the user convenience of simple
  text searching with the expressive refinements of powerful query
  languages. The database query user interface includes a query text string
  input from a user including one or more terms of a chunk expression
  language format. The database query user interface further includes a
  syntactical prompt for constructing a multi-element chunk expression
  language database query that is syntactically correct and complete and
  includes the text string input from the user. For example, the
  syntactical prompt is selected from the database based upon a weighted
  analysis of database information relating to database elements included
  in the text string input from the user. A database query formed according
  to the present user interface may then be persisted or stored as a
  database query object.

French Abstract
  L'invention concerne une interface-utilisateur d'interrogation de base de
  donnees qui combine la facilite de la recherche en texte simple aux
  subtilites des langages d'interrogation puissants. Cette
  interface-utilisateur d'interrogation de base de donnees utilise une
  chaine de caracteres d'interrogation saisie par l'utilisateur, comprenant
  un ou plusieurs termes d'un format de langage d'expression segmente.

L'interface-utilisateur d'interrogation de base de donnees utilise
egalement une invite syntaxique permettant de former une interrogation de
base de donnees en langage d'expression segmente multi-element
syntaxiquement correcte et complete et utilise la chaine de caracteres
saisie par l'utilisateur. Par exemple, l'invite syntaxique est
selectionnee dans la base de donnees sur la base d'une analyse ponderee
des informations de la base de donnees concernant les elements de la base
de donnees inclus dans la chaine de caracteres saisie par l'utilisateur.
Une interrogation de base de donnees formee au moyen de la presente
interface-utilisateur peut ensuite etre reutilisee ou stockee sous forme
d'objet d'interrogation de base de donnees.

Detailed Description
... results of a query.
  [0017] Fig. 1 1 is a flow diagram of an automated **query** re-write method
  to create queries automatically by making a simpler query and adding
  exceptions...

...modified query incorporating a
suggested rewrite.

  [0021] Fig. 15 is an illustration of a grammatical **query** autocomplete
  (GQA) user interface.

3
[0022] Fig. 16 is a flow diagram of a GQA is a grammatical **query**
autocomplete (GQA).

  [0023] Fig. 17 is a flow diagram of a **query   weighting** and **sorting**
  method for **weighting** and **sorting** the results returned from internal
  queries .

Detailed Description of Preferred Embodiments
[0024] A conventional database system or database includes a
collection of tables with record entries. **Queries** of the database are
typically made using a **query** specification language, sometimes
referred to as a data manipulation language, such as SQL. In
addition, a full-text **search** engine can **find** records that contain
text
strings. A variety of commercially available databases are available,
including Microsoft SQL available from Microsoft Corporation. The
 term database is used herein to refer generally to any "property store"
that includes objects or **files** with searchable **properties** .

[0025] Fig. 1 is a flow chart of a simplified representation of a prior
art
database **query** sequence 100. In a step 102, a user determines that
he or she wants particular...

01173034   **Image available**
RELATIONSHIP VIEW
VISUALISATION DE RELATIONS
Patent Applicant/Assignee:
  MICROSOFT CORPORATION, One Microsoft Way, Redmond, WA 98052, US, US
     (Residence), US (Nationality)
Inventor(s):
  DRUCKER Steven M, 22555 W. Lake Sammamish Parkway, SE, Bellevue, WA 98008
     , US,
  WONG Curtis G, 301 109th Avenue, SE, Bellevue, WA 98004, US,
  GLATZER Asta L, 7417 Old Redmond Road, Redmond, WA 98052, US,
Legal Representative:
  AMIN Himanshu S (et al) (agent), Amin & Turocy, LLP, 1900 E. 9th Street,
     24th Floor, National City Center, Cleveland, OH 44114, US,
Patent and Priority Information (Country, Number, Date):
  Patent: WO 200495237 A2 20041104 (WO 0495237)
  Application: WO 2004US9190 20040326 (PCT/WO US04009190)
  Priority Application: US 2003420414 20030422
Designated States:
(All protection types applied unless otherwise stated - for applications
2004+)
  AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM
  DZ EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC
  LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NA NI NO NZ OM PG PH PL PT RO
  RU SC SD SE SG SK SL SY TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW
  (EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PL PT RO
  SE SI SK TR
  (OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
  (AP) BW GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
  (EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: **G06F**
Publication Language: English
Filing Language: English
Fulltext Availability:
  Detailed Description
  Claims
Fulltext Word Count: 11548

English Abstract
  The present invention provides a unique method and user interface that
  facilitates accessing and browsing objects in which a user begins with a
  center object (e.g., one or a few focal objects) displayed on a screen
  and related objects are populated on the screen as well. The related
  objects can be further organized into clusters whereby each cluster or
  grouping of objects expands on a particular attribute of the center
  object. The attributes correspond to metadata. Thus, the objects are
  populated based upon the metadata of the center object. According to one
  aspect, the user can access one or more specific objects having a
  plurality of attributes and then relax at least one of the attributes to
  see what other objects share at least one attribute with the center
  object. According to another aspect, the object having the closest match
  to a search request can be centrally displayed with other close matches
  arranged by their respective metadata.

French Abstract
  L'invention concerne un procede unique et une interface utilisateur
  facilitant l'acces et l'exploration d'objets, procede selon lequel un
  utilisateur commence par un objet central (par exemple, un ou quelques
  objets focaux) affiches sur un ecran, des objets apparentes peuplant
  egalement l'ecran. Les objets apparentes peuvent etre ulterieurement
  organises en grappes, chaque grappe ou groupement d'objets s'etendant sur

un attribut particulier de l'objet central. Les attributs correspondent a des meta-donnees. De cette facon, les objets sont peuples sur la base des meta-donnees de l'objet central. Conformement a un aspect de l'invention, l'utilisateur peut acceder a un ou plusieurs objets specifiques ayant une pluralite d'attributs, puis relacher au moins l'un des attributs pour voir quel autre objet partage au moins un attribut avec l'objet central. Conformement a un autre aspect, l'objet correspondant le plus a une demande de recherche peut etre affiche au centre, avec d'autres correspondances proches, agencees par leurs meta-donnees respectives

Legal Status (Type, Date, Text)
Publication   20041104 A2 Without international search report and to be republished upon receipt of that report.
Examination   20050922 Request for preliminary examination prior to expiration of applicable time limit under Rule 54bis.1(a)

Main International Patent Class: **G06F**
Fulltext Availability:
  Detailed Description

Detailed Description
...   0 has metadata associated therewith and can be received for example by a user-based **search** request mechanism. Other mechanisms can also be employed to receive the first object.
One approach to the process 300 involves performing a non-specific **search** request using one or more attributes (metadata) such as when the desired object is not known. For example, when a user would like to **find** a particular book title written by Stephen King or one of his other pseudonyms published...

...the year it was published, a user can enter
1
one or more non-specific **search** **terms** in **order** to **retrieve** an object somewhat related to or in the neighborhood of the desired object (e.g...

...user. At 330, a plurality of additional objects (e.g., book titles, movies, websites, news **articles** , etc.) having respective **metadata** associated therewith. The respective inetadata of the additional objects are at least in part related...

...Stand" book cover (e.g., first object). Metadata associated with the first object can be **weighted** to determine the strength of correlation between the first object and other objects selected for clustering. The **weight** of each metadata associated with the first object can be determined based at least in part upon user input (e.g., via a user-based **search** request).

In one aspect of the present invention, objects having the strongest correlation to the...

01121949   **Image available**
METHOD AND APPARATUS FOR PROCESSING FILES UTILIZING A CONCEPT OF WEIGHT SO AS TO VISUALLY REPRESENT THE FILES IN TERMS OF WHETHER THE WEIGHT THEREOF IS HEAVY OR LIGHT METHOD AND APPARATUS FOR PROCESSING FILES UTILIZING A CONCEPT OF WEIGHT SO AS TO VISUALLY REPRESENT THE FILES IN TERMS OF WHETHER THE WEIGHT THEREOF IS HEAVY OR LIGHT
METHODE ET APPAREIL DE TRAITEMENT DE FICHIERS UTILISANT UN CONCEPT A BASE DE POIDS POUR LES REPRESENTER VISUELLEMENT SELON LEUR POIDS
Patent Applicant/Assignee:
  SONY COMPUTER ENTERTAINMENT INC, 2-6-21, Minami-Aoyama, Minato-ku, Tokyo 107-0062, JP, JP (Residence), JP (Nationality)
Inventor(s):
  WADA Shinya, c/o SONY COMPUTER ENTERTAINMENT INC., 2-6-21, Minami-Aoyama, Minato-ku, Tokyo 107-0062, JP,
Legal Representative:
  MORISHITA Sakaki (agent), 2-11-12, Ebisu-Nishi, Shibuya-ku, Tokyo 150-0021, JP,
Patent and Priority Information (Country, Number, Date):
  Patent: WO 200444728 A2-A3 20040527 (WO 0444728)
  Application: WO 2003JP14286 20031110 (PCT/WO JP03014286)
  Priority Application: JP 2002328853 20021112
Designated States:
(Protection type is "patent" unless otherwise stated - for applications prior to 2004)
  AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
  EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS KE KG KP KR KZ LC LK LR
  LS LT LU LV MA MD MG MK MN MW MX MZ NI NO NZ OM PG PH PL PT RO RU SC SD
  SE SG SK SL SY TJ TM TN TR TT TZ UA UG UZ VC VN YU ZA ZM ZW
  (EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT RO SE SI SK TR
  (OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
  (AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
  (EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: **G06F-003/033**
Publication Language: English
Filing Language: English
Fulltext Availability:
  Detailed Description
  Claims
Fulltext Word Count: 12179

English Abstract
  On a screen a file processing apparatus displays values of an attribute related to a plurality of files using a concept of weight. The weight of each file is represented by spherical objects submerged into the water and displayed on the screen. For example, a first spherical object represents a file whose data size is large, and is sunk near bottom. And a second spherical object represents a file whose data size is light and is floating near the water surface.

French Abstract
  L'invention porte sur un appareil de traitement de fichiers presentant sur un ecran les valeurs d'attributs de differents fichiers en utilisant un concept a base de poids. A chaque fichier est attribue un poids represente par un objet spherique immerge dans de l'eau, apparaissant sur l'ecran. Par exemple un premier objet spherique representant un gros fichier de donnees flotte au voisinage du fond, tandis qu'un deuxieme objet spherique representant un petit fichier de donnees flotte au voisinage de la surface.

Legal Status (Type, Date, Text)
Publication   20040527 A2 Without international search report and to be
                       republished upon receipt of that report.
Examination   20040812 Request for preliminary examination prior to end of
                       19th month from priority date
Search Rpt   20041216 Late publication of international search report
Republication 20041216 A3 With international search report.

Main International Patent Class: **G06F-003/033**
Fulltext Availability:
  Detailed Description

Detailed Description
... a file are
  displayed simply as a character string on the screen, the
  user may **find** it difficult to read them when selecting a
  desired file. This may also develop into...

...files at the
  respective display positions on a screen, and expressing
  visually comparison of the **weights** of the objects via
  another object that symbolizes **weight** measurement.

  "Another object that symbolizes **weight** measurement" is
  a character that is used to display on the screen a
  comparison of **weights** between files or a measurement of a
  total **weight** of a plurality of files. Such an object used
  for instance is a weighing device...

...This method includes: acquiring values of a
  predetermined attribute for a plurality of files, in **order**
  to represent the values of a predetermined **attribute** for
  intended **files** by using a concept of **weight** ; setting a
  temporary sequence for the plurality of files; determining,
  based on the temporary sequence, a temporary display
  position of a predetermined object that symbolically
  represents the files in **terms** of whether the **weight** thereof
  is heavy or light; displaying an object that corresponds to
  the plurality of files, at the temporary display position on
  a screen; comparing the values of a predetermined **attribute**
  between adjacent **files** in the temporary sequence; updating
  the display position based on a comparison result obtained
  from the comparing; and representing visually the **weight**
  thereof by varying display contents according to the
  updating.

  A "temporary sequence" may be a temporary **order** of
  arrangement for convenience, sake, for instance, when
  displaying on the screen a plurality of...
...still an initial display state, and
  the values of an attribute using the concept of **weight** are
  not yet represented.

  "Adjacent files" are not necessarily strictly adjacent
  to each other in...

16/5,K/15   (Item 15 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2005 WIPO/Univentio. All rts. reserv.

01082684   **Image available**
SYSTEM AND METHOD OF CREATING AND USING COMPACT LINGUISTIC DATA
SYSTEME ET PROCEDE DE GENERATION ET D'UTILISATION DE DONNEES LINGUISTIQUES
    COMPACTEES
Patent Applicant/Assignee:
  2012244 ONTARIO INC, 295 Phillip Street, Waterloo, Ontario N2L 3W8, CA,
    CA (Residence), CA (Nationality), (For all designated states except:
    US)
Patent Applicant/Inventor:
  FUX Vadim, 575 Rustic Drive, Waterloo, Ontario N2K 2A4, CA, CA
    (Residence), IL (Nationality), (Designated only for: US)
  ELIZAROV Michael G, 465 Kingscourt Drive, Unit 3, Waterloo, Ontario N2K
    3R5, CA, CA (Residence), IL (Nationality), (Designated only for: US)
Legal Representative:
  PATHIYAL Krishna K (et al) (agent), Research In Motion Limited, 295
    Phillip Street, Waterloo, Ontario N2L 3W8, CA,
Patent and Priority Information (Country, Number, Date):
  Patent: WO 200406122 A2 20040115 (WO 0406122)
  Application: WO 2003CA1023 20030703 (PCT/WO CA03001023)
  Priority Application: US 2002393903 20020703; US 2002289656 20021107; CA
    2411227 20021107
Designated States:
(Protection type is "patent" unless otherwise stated - for applications
prior to 2004)
  AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ
  EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR
  LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SC SD SE SG
  SK SL TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW
  (EP) AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL PT RO SE
  SI SK TR
  (OA) BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG
  (AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW
  (EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: **G06F-017/27**
Publication Language: English
Filing Language: English
Fulltext Availability:
  Detailed Description
  Claims
Fulltext Word Count: 8850

English Abstract

French Abstract

Legal Status (Type, Date, Text)
Publication 20040115 A2 Without international search report and to be
                        republished upon receipt of that report.
Declaration   20040422 Late publication under Article 17.2a
Republication 20040422 A2 With declaration under Article 17(2)(a); without
                        abstract; title not checked by the International
                        Searching Authority.

Main International Patent Class: **G06F-017/27**
Fulltext Availability:
  Detailed Description

Detailed Description
... analyzer 202 illustrated in Fig. 6.

The absolute frequency of a certain group of words **found** in the corpus 200 may alternatively be modified by separating this group to a different file lo and assigning a custom **weight** to this file. This group may consist of words which are domain specific, such as...

...absolute value of the frequencies for this group of words will be modified using the **weight** assigned to the group, so that this group of **words** will have frequencies that are different they would have otherwise had.

Fig. 3 is flowchart...

...comprises a corpus. The filtering method is the first step in calculating the frequency of **words** in the corpus.

The method begins with the step 300 of reading the contents a...

...of text,from the file according to user preferences, which may be stored in a **properties** **file** . The user preferences specify regular expressions which are applied to the text in **order** to substitute invalid or unwanted characters. For example, a user may not want street names included in the **word** 8 list, or an Italian user may want to replace "e"' followed by a non...

```
Set     Items    Description
S1     7234797    SEARCH?? OR SEARCHING OR FIND? ? OR FINDING OR FOUND OR RE-
                  TRIEVE? ? OR RETRIEVING OR RETRIEVAL OR QUERY OR QUERIES OR Q-
                  UERYING
S2     3363549    KEYWORD? ? OR PHRASE? ? OR TERM? ? OR WORD? ?
S3     9853729    ATTRIBUTE? ? OR CHARACTERISTIC? ? OR PROPERTY OR PROPERTIES
                  OR METADATA
S4     3955618    SCALE? '? OR SCALING OR SCORE? ? OR SCORING OR WEIGHT?? OR -
                  WEIGHTING
S5     4489378    ORDER?? OR ORDERING OR SORT?? OR SORTING
S6        4114    S1 AND S2 AND S3 AND S4 AND S5
S7     1985265    FILE? ? OR DOCUMENT? ? OR ARTICLE? ? OR WEBPAGE? ? OR WEBS-
                  ITE?
S8       15580    S7 (3N) S3
S9          13    S1 AND S2 AND S8 AND S4 AND S5
S10         10    S9 NOT PY>2001
S11          9    RD   (unique items)
S12    1098355    WEB OR WEBPAGE? ? OR WEBSITE? ? OR ONLINE OR ON()LINE OR I-
                  NTERNET? ? OR INTRANET? EXTRANET? ? OR WWW  OR WORLDWIDE()WEB
S13        106    S6 AND S12
S14         21    S1 (10N) S2 (10N) S12 AND S3 AND S4 AND S5
S15         20    S14 NOT S11
S16         11    S15 NOT PY>2001
S17          8    RD   (unique items)
File   8:Ei Compendex(R) 1970-2006/Jan W1
          (c) 2006 Elsevier Eng.  Info. Inc.
File  35:Dissertation Abs Online 1861-2005/Dec
          (c) 2005 ProQuest Info&Learning
File  65:Inside Conferences 1993-2006/Jan W2
          (c) 2006 BLDSC all rts. reserv.
File   2:INSPEC 1898-2006/Dec W3
          (c) 2006 Institution of Electrical Engineers
File  94:JICST-EPlus 1985-2006/Oct W5
          (c)2006 Japan Science and Tech Corp(JST)
File 111:TGG Natl.Newspaper Index(SM) 1979-2006/Jan 04
          (c) 2006 The Gale Group
File   6:NTIS 1964-2006/Dec W3
          (c) 2006 NTIS, Intl Cpyrght All Rights Res
File 144:Pascal 1973-2006/Dec W3
          (c) 2006 INIST/CNRS
File 434:SciSearch(R) Cited Ref Sci 1974-1989/Dec
          (c) 1998 Inst for Sci Info
File  34:SciSearch(R) Cited Ref Sci 1990-2006/Jan W1
          (c) 2006 Inst for Sci Info
File  62:SPIN(R) 1975-2006/Oct W5
          (c) 2006 American Institute of Physics
File  99:Wilson Appl. Sci & Tech Abs 1983-2005/Nov
          (c) 2006 The HW Wilson Co.
File  95:TEME-Technology & Management 1989-2006/Jan W2
          (c) 2006 FIZ TECHNIK
File  56:Computer and Information Systems Abstracts 1966-2005/Dec
          (c) 2005 CSA.
File  57:Electronics & Communications Abstracts 1966-2005/Dec
          (c) 2005 CSA.
```

940984 ORDER NO: AAD87-02080
A COMPOSITE MEASURE FOR WEIGHTING **DATABASES IN DEFENSE, ENGINEERING, AND
SCIENCE (MULTIDATABASE** SEARCHING **, ONLINE** SEARCHING **, DIALINDEX)**
    Author: RITTENHOUSE, ROBERT JOHN
    Degree: PH.D.
    Year: 1986
    Corporate Source/Institution: CASE WESTERN RESERVE UNIVERSITY (0042)
    Source: VOLUME 47/10-A OF DISSERTATION ABSTRACTS INTERNATIONAL.
        PAGE 3598. 298 PAGES
    Descriptors: INFORMATION SCIENCE
    Descriptor Codes: 0723

The primary problem of this dissertation is to propose a composite
measure as a technique for measuring the relevancy of databases. The
databases are characterized as single units by the measure of closeness,
C(,M), values. The measure of closeness consists of two **weighted** factors:
(1) a relevance factor, and (2) a descriptive factor. The relevance factor
is the sum of the recall and precision ratios. The descriptive factor is
the sum of the **weighted** **properties** of each **file** as follows: (1)
subject coverage, (2) thesaurus strength, (3) technical level, (4) subject
coding, and (5) length of years **searched** retrospectively.
    Two experiments were conducted to test if the measure of closeness
may be utilized to select the relevant databases in DIALINDEX **searches** in
the general areas of defense, engineering, and science. Databases from
Dialog Information Services, Inc., Defense Logistics Studies Information
Exchange, Defense Technical Information Center, Mead Data Central Nexis,
NASA/RECON, and DOE/RECON were also used. **Searches** were conducted in
seven sample topics: (1) composites, (2) missiles, (3) rockets, (4) sonar,
(5) torpedoes, (6) underwater acoustics, and (7) underwater weapons.
    For each of the seven topics, online **searches** were performed on
a group of databases. These databases, ranked according to C(,M) values,
were compared with their corresponding databases ranked by retrievals from
DIALINDEX, a Dialog multidatabase file. The first experiment compared six
randomly selected Dialog files and Dialog files subjectively selected for
their expected higher relevance to the topics. While randomly selected
files **retrieved** some relevant citations, these files generally did not
contain many relevant citations. The second experiment compared the
DIALINDEX method and the measure of closeness, C(,M), technique.
    Mann-Whitney two rank and Spearman Rho rank correlation tests
failed to indicate conclusively that the DIALINDEX method is different from
use of the **weighted** measure of closeness alone. The tests did indicate
DIALINDEX **term** frequency retrievals appear to result in ranking relevant
databases. Possible artificial intelligence designs may further enhance the
future modelling of **weighting** schemes for more effective multivendor and
multidatabase online **search** techniques.
    Only unclassified **terms**, titles and/or abstracts were discussed
in **order** to conform to U.S. national security requirements.

04182617   INSPEC Abstract Number: C88047289
  Title: Probabilistic design principles for conventional and full-text
retrieval **systems**
  Author(s): Maron, M.E.
  Author Affiliation: Sch. of Libr. & Inf. Studies, California Univ.,
Berkeley, CA, USA
  Journal: Information Processing & Management   vol.24, no.3   p.249-55
  Publication Date: 1988 Country of Publication: UK
  CODEN: IPMADK ISSN: 0306-4573
  U.S. Copyright Clearance Center Code: 0306-4573/88/$3.00+.00
  Language: English   Document Type: Journal Paper (JP)
  Treatment: Practical (P)
  Abstract: In **order** for conventionally designed commercial document
retrieval systems to perform perfectly, the following two (logical)
conditions must be satisfied for every **search** : (1) There exists a
document property (of combination or properties) that belongs to those
(and only those) documents that are relevant. (2) That property (or
combination of properties) can be correctly guessed by the **searcher** . In
general, the first assumption is false, and the second is impossible to
satisfy; hence no conventional IR system can perform at a maximum level of
effectiveness. However, different design principles can lead to improved
performance. The article presents a view of the document **retrieval**
 problem that shows that since the relationship between **document**
 properties (whether they be humanly assigned index **terms** or **words**
that occur in the running text) and relevance is at best probabilistic, one
should approach the design problem using probabilistic principles. It turns
out that a front end designed to permit searchers to attach
probabilistically interpreted **weights** to their **query terms** could be
adapted for conventional IR systems. Such an enhancement could lead to
improved performance. (37 Refs)
  Subfile: C
  Descriptors: information **retrieval**
  Identifiers: conventional; full-text **retrieval** systems; document
retrieval systems; **document property** ; IR system; document **retrieval**
problem; probabilistic principles; front end; probabilistically interpreted
 weights ; **query terms**
  Class Codes: C7250 (Information storage and retrieval)

07690504   INSPEC Abstract Number: C2000-10-7210N-027
   **Title: ROC performance evaluation of Web-based bibliographic navigator
using extended association rules**
   Author(s): Kawahara, M.; Kawano, H.
   Author Affiliation: Data Process. Center, Kyoto Univ., Japan
   Conference Title: 5th International Computer Science Conference ICSC'99.
Proceedings (Lecture Notes in Computer Science Vol. 1749)   p.216-25
   Editor(s): Hui, L.C.-K.; Lee, D.L.
   Publisher: Springer-Verlag, Berlin, Germany
   Publication Date: 1999 Country of Publication: Germany   xx+518 pp.
   ISBN: 3 540 66903 5   Material Identity Number: XX-1999-03571
   Conference Title: Proceedings of ICSC'99: 5th International Computer
Science Conference
   Conference Date: 13-15 Dec. 1999   Conference Location: Hong Kong, China
   Language: English   Document Type: Conference Paper (PA)
   Treatment: Practical (P)
   Abstract: It is very effective for **search** users to provide meaningful
 keywords which are derived by text mining algorithm. We are developing
our **search** engine "Mondou" using **weighted** association rules, as a **Web**
-based intelligent database navigation system. In this paper, we focus on
the computing cost to derive appropriate keywords, we carefully determine
system parameters, such as Minsup and Mincon f threshold values. In **order**
 to evaluate the performance and **characteristics** of derived rules, we use
the techniques of ROC graph. We propose the ROC analytical model of our
search system, and we evaluate the performance of **weighted** association
rules by the ROC convex hull method. Especially, we try to specify the
optimal threshold values to derive effective rules from INSPEC database,
which is a huge bibliographic database. (6 Refs)
   Subfile: C
   Descriptors: bibliographic systems; deductive databases; information
resources; optimisation; search engines; software performance evaluation;
text analysis; very large databases
   Identifiers: ROC performance evaluation; Web-based bibliographic
navigator; extended association rules; meaningful keywords; text mining
algorithm; search engine; Mondou; **weighted** association rules; Web-based
intelligent database navigation system; Minsup threshold value; Mincon f
threshold value; ROC graph; ROC convex hull method; optimal threshold
values; INSPEC database; bibliographic database
   Class Codes: C7210N (Information networks); C7250C (Bibliographic
retrieval systems); C7240 (Information analysis and indexing); C6160K (
Deductive databases); C1180 (Optimisation techniques)

05042893   JICST ACCESSION NUMBER: 01A1008142   FILE SEGMENT: JICST-E
**An Interactive Method for Supporting WWW Retrieval Based on Adjustable
    Ranking of Documents.**
KINOSHITA ATSUFUMI (1); NAKAGAWA KOKORO (1); TAKADA YOSHIAKI (1); SEKI
    HIROYUKI (1)
(1) Advanced Inst. Sci. and Technol., Nara
ABSTRACT: We propose a method for supporting **WWW** document **retrieval**
    which allows a user to adjust a ranking of documents. Many **keyword**
    -based **search** services are available and these services provide a
    user with a ranked list of documents, arranged in a descending **order**
    of relevancy to input keywords. Various **scoring** methods to rank
    documents have been proposed; however, a highly-ranked document is not
    always a desirable one for the user. This difference puts stress on the
    user because not only is it difficult to find desirable documents, but
    also no means to adjust the ranking is provided. To overcome the
    problem, we propose three methods for allowing a user to directly
    adjust ranking of documents. A retrieval system based on the proposed
    method has been implemented and experimental results on 6 human
    subjects are presented. (author abst.)
DESCRIPTORS: document retrieval; query processing; WWW(communication);
    keyword; contraction(mathematics); similarity; statistical estimation;
    interactive processing; recall precision; performance evaluation
IDENTIFIERS: feature vector; degree of association; ranking
BROADER DESCRIPTORS: information retrieval; retrieval; information
    processing; treatment; information system; computer application system;
    system; vocabulary; **property** ; estimation; statistical decision;
    decision; statistical method; efficiency; evaluation
CLASSIFICATION CODE(S): AC06020S; JD03030U; JE15050M

04008960   JICST ACCESSION NUMBER: 99A0354360   FILE SEGMENT: JICST-E
Studies on User Adaptive Browsing System on Internet.
IMANAKA TAKESHI (1); OZAWA JUN (1); MIURA KOJI (1); MATSUURA SATOSHI (1)
(1) Matsushita Electr. Ind. Co., Ltd., Cent. Res. Lab.
ABSTRACT: This paper describes two of our studies on user adaptive browsing
    system on **Internet** . One is a Kansei **retrieval** system which enables
    users to **retrieve** database by Kansei **words** like "cheerful", "happy"
    and so on. In this system, we introduce Kansei model to define
    individuality in users' feelings about web pages on WWW (World Wide
    Web). By making use of the model, the system can adjust database to
    each user's Kansei in retrieving. The other is a browsing system with
    user adaptive index which is constructed by **weighted** keywords in HTML
    text. The **weight** of keywords in the index is computed by our proposed
    event driven rules called XECA (eXtended Event Condition Action) rules.
    XECA rules are defined as an extension of conventional ECA rules to
    deal with time constraints. In our system, the index involves user's
    preferences which are extracted automatically depending on how the user
    has browsed pages(i.e. **order** , time). As a result, the user can find
    favorite web pages easily by using the index. (author abst.)
DESCRIPTORS: information retrieval; machine learning; sensibility; adaptive
    system; browsing; keyword; index; **weighting** ; constraint
    condition(restriction); database
BROADER DESCRIPTORS: retrieval; learning; sensitivity; **property** ; system;
    reading(library); action and behavior; vocabulary; condition
CLASSIFICATION CODE(S): JE15050M; JE08000Z

02773495   JICST ACCESSION NUMBER: 96A0736608   FILE SEGMENT: JICST-E
Implementation and Evaluation of WWW Search System RCAAU.
NISHIMURA HIDEKI (1); KAWANO HIROYUKI (2); HASEGAWA TOSHIHARU (2)
(1) Sharp Corp.; (2) Kyoto Univ., Graduate School
ABSTRACT: We have been developing **WWW   search** system with the several
     advanced functions, such as **keywords** focusing by a data mining
     techniques and network **characteristics** evaluation. In this paper, we
     explain implementation of **WWW** robot which collects data from **Web**
     servers and stores several **attributes** into database. Then, based on
     http access log, we analyze **keywords** of **queries** and its embedded
     tendency. Moreover, in **order** to investigate throughly how effective
     our functions are for users, we evaluate the quality of **keywords**
     derived by **weighted** association rule. (author abst.)
DESCRIPTORS: database; information retrieval; correlation function;
     knowledge acquisition; keyword; recall precision; data collection;
     search problem
BROADER DESCRIPTORS: retrieval; function(mathematics); mapping(mathematics)
     ; acquisition; vocabulary; efficiency; information collection;
     collection; information processing; treatment; problem
CLASSIFICATION CODE(S): JD03030U; AC06020S

DIALOG(R)File 350:Derwent WPIX
(c) 2006 Thomson Derwent. All rts. reserv.

015268212   **Image available**
WPI Acc No: 2003-329141/200331
XRPX Acc No: N03-263277
  **Web page search method e.g. for mathematics web page involves calculating
  overall matching score for ordering selected web page, based on
  determined criterion matching score and associated scaling factor**
Patent Assignee: MASTERS G S (MAST-I)
Inventor: **MASTERS G S**
Number of Countries: 001 Number of Patents: 001
Patent Family:
Patent No	Kind	Date	Applicat No	Kind	Date	Week	
US 20020198875	A1	20021226	US 2001885902	A	20010620	200331	B

Priority Applications (No Type Date): US 2001885902 A 20010620
Patent Details:
Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
US 20020198875	A1		14	G06F-007/00	

Abstract (Basic): US 20020198875 A1
    NOVELTY - A search criterion associated with a keyword match
between a keyword entry and identified web pages, is established based
on the attribute of the web pages. A criterion matching score for the
web pages is determined and a scaling factor is associated with the
search criteria for calculating an overall matching score based on
which the selected web page is ordered.
    DETAILED DESCRIPTION - INDEPENDENT CLAIMS are included for the
following:
    (1) web page search engine; and
    (2) web page search computer system.
    USE - For searching database storing web page comprising
information about particular subject e.g. mathematics, English
literacy, other languages, computer science, etc., by high school
student using web page search computer system (claimed).
    ADVANTAGE - The user is enabled to conduct search of web page that
simultaneously takes account of keyword matching and web page
attributes. The user is enabled to easily vary and adjust the relative
weighting of the search criteria for optimizing search result.
    DESCRIPTION OF DRAWING(S) - The figure shows a flowchart explaining
the web page search process.
    pp; 14 DwgNo 1/9
Title Terms: WEB; PAGE; SEARCH; METHOD; MATHEMATICAL; WEB; PAGE; CALCULATE;
  OVERALL; MATCH; SCORE; ORDER; SELECT; WEB; PAGE; BASED; DETERMINE;
  CRITERIA; MATCH; SCORE; ASSOCIATE; SCALE; FACTOR
Derwent Class: T01
International Patent Class (Main): **G06F-007/00**
File Segment: EPI